基因表現連續分析法（SAGE）及其在植物研究之應用

農業試驗所　王強生　廖仁盟

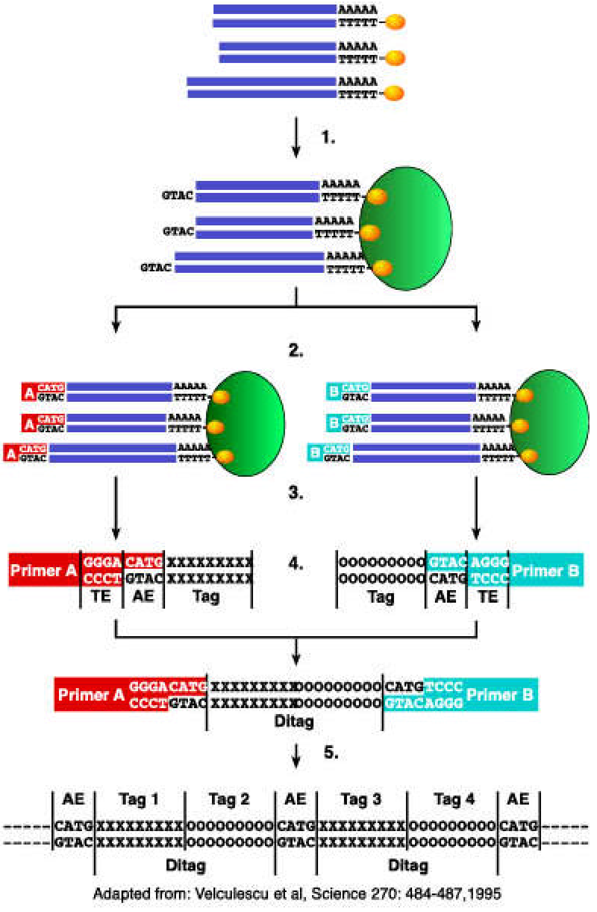
近來由於人類、阿拉伯芥及水稻等各物種基因組計畫之快速發展，基因序列之資訊正以史無前例的速度迅速累積中。然而人類對於基因功能之瞭解程度卻遠遠落後。科學家長久以來便對差異性的基因頗感興趣，一直希望能依據差異選殖基因(cloning by difference)，但是礙於技術層面，往往未能如願。這些年隨著分子生物學的進展及技術層面的提昇，選殖具有差異性的基因已經不再是夢想。

傳統上應用於分析基因表現的技術有：北方轉漬、RNA 點轉漬及 RT-PCR 方法，但每次僅能針對有限基因進行評估分析。其他如：扣除雜交法(subtractive hybridization)、比較 EST 分析法(comparative EST analysis)是用來偵測基因表現組群(profiles)的變化。而基因差異篩選法(differential screening)及基因差異展示法(differential display, DD)則應用於鑑定二個具有差異表現的樣品基因並予分離。然而，這些方法並無法大規模有效、準確的偵測基因差異表現，每次僅能分析有限數目之轉錄訊息，更無法提供基因表現量的資料。此外，這些方法一般也僅能偵測大量表現的 mRNA。

最近，由 Velculescu 等人(1995)所發展的「基因表現連續分析法」 (serial analysis of gene expression, SAGE)可以準確的在同一時間內針對數千個基因進行基因表現量的分析，因此 SAGE 已成為一研究後基因組階段(post-genomic era)基因表現的有力工具。茲將 SAGE 法之原理簡介如下：

此法乃利用 *Nla* III 限制酵素，將最靠近 cDNA 3’ 端切位的一個 9-11 個鹼基對的短片段 DNA (即標記)，利用連續的 Linker-Ligations 及限制酵素切割的方法擷取，進而利用 PCR 擴增標記並予選殖、定序，而每一標記在序列出現頻度可直接反應每一 mRNA 在組織中表現量的多寡(abundance)。如果標記可自轉錄訊息中細微位置萃取而得，則 9 個鹼基對所包含的訊息(4^9)已遠較植物中所有基因數為多。據估計，阿拉伯芥中約有 15,000 – 33,000 個基因，因此可以確定每一個植物基因均可有一獨特的 SAGE 標記來區分。

SAGE 分析法之操作步驟(見圖一)：



圖一、基因表現連續分析法 SAGE 之流程圖（引用自 SAGE 網站http://www.sagenet.org/）。

一、Total RNA 萃取及 mRNA 純化

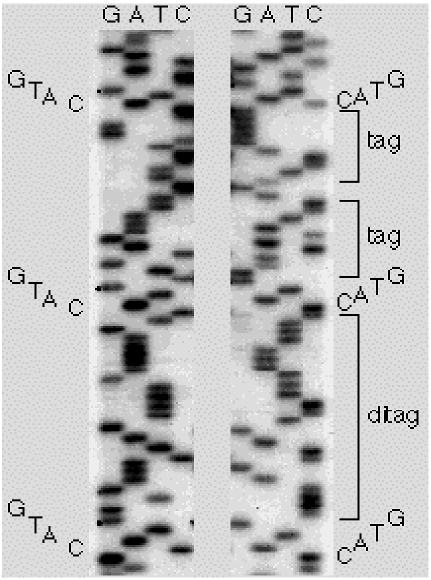
1. 利用 phenol/chloroform 法(或其他方法)進行 total RNA 的萃取，欲進行 SAGE 所需的 total RNA 最少需 1 mg (假設 mRNA 的含量為 5 %)以上。
2. 採用由 Dynal® Co. 所生產之 Dynabeads® Oligo (dT)$_{25}$ 試劑組並依照使用手冊進行 polyA RNA 的分離，待定量後，取 5 μg polyA RNA 進行第一股 cDNA 的合成。
3. 在微量離心管中依序加入 first stranded buffer、dNTP mixture、biotinylated oligo dT primer 與 5 μg polyA RNA，於 65°C水浴器中 10 分鐘以打開 RNA 的二級結構，置於冰上 3 分鐘，加入 MMLV reverse transcriptase 於 37°C進行第一股 cDNA 的合成。
4. 將步驟 3 的雙股 cDNA:RNA 以 RNase H 及 DNA polymerase I 進行第二股 cDNA 的合成，合成的雙股 cDNA 經定量後取適當量進行限制酵素切割。

二、基因表現連續分析法

1. 使用可能於 cDNA 中切位(CATG)較多的限制酵素(frequent cut enzyme, 如 *Nla* III)(即 AE)進行切割，切割後的 cDNA 以 streptavidin beads 回收具有 polyA 的 3'端 cDNA。
2. 將回收之 cDNA 分成二部分，分別接上特殊的 Linker A 及 Linker B (此 linker 上均具有 Type IIS 的限制酵素切位)，回收之後分別再以 Type IIS 限制酵素(如 *Fok* I)(TE)切割。
3. 切割完成之 cDNA 以 Klenow fragment 將末端補齊(fill in)為齊平端(blunt

end)，經過 phenol/chloroform 萃取與酒精沉澱後，將二管 cDNA 集中成一管進行 ligation。

4. 粘接完成的 cDNA 以 phenol/chloroform 萃取與經酒精沉澱後再以第一次使用的限制酵素切割(如 *Nla* III)，經切割後的 cDNA 上則具有雙標記(ditag)，片段大小為 26-30 個鹼基(bases)。
5. 由於具有雙標記的 cDNA 片段太小，難以進行選殖，故將上述的 cDNA 進行黏接作用，完成後進行片段分離(size fractionation)，回收大於 200 個鹼基的片段，此時片段內含有多個雙標記。
6. 將回收的片段選殖至載體上並進行轉形作用，挑選含有選殖片段的殖系進行定序反應(見圖二)後，利用 SAGE 軟體進行序列分析，找出基因表現的標記，並利用 BLAST analysis 程式蒐尋基因庫。SAGE 標記之序列經基因庫比對後可能有三種結果：a.為已知基因，即相匹配(match)；b.與已知 ESTs (expressed sequence tags)或不知名的 mRNA 相似；c.目前為一未經確認功能的 mRNA。



圖二、經 SAGE 分析產生的單標記及雙標記產物之 DNA 定序分析
（引用自 SAGE 網站 http://www.sagenet.org/）。

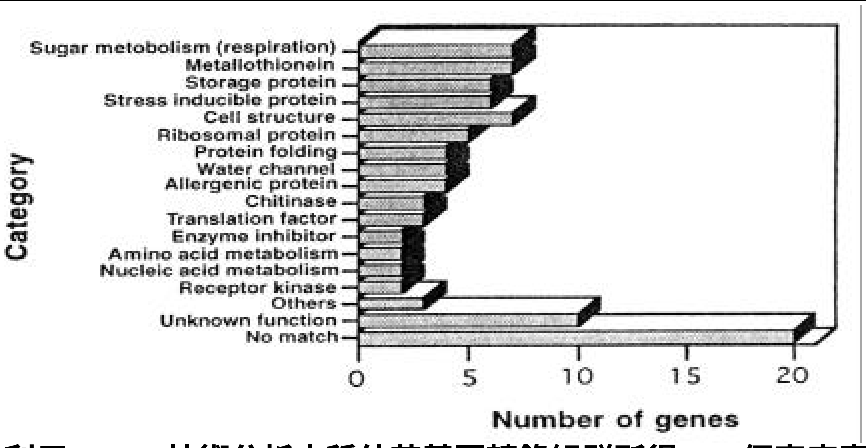
目前 SAGE 分析法已成功的應用於鑑定酵母菌轉錄群組(transcript profiling)，及區分人類正常細胞與癌細胞的差異轉錄訊息。從比較癌細胞與 p53 誘導的細胞老死(apoptosis)結果顯示，SAGE 分析法能極有效的解析受轉錄因子調控的連續遺傳步驟。以下僅以 Matsumura 等(1999)利用 SAGE 技術研究水稻幼苗基因表現為例，說明 SAGE 在植物基因表現分析之應用情形及其可能展望：

自從 SAGE 技術發展以來，絕大部分集中於人類與動物方面的研究，將此技術成功應用於植物上者，首推 Matsumura 等(1999)的試驗，嘗試利用此一新技術建立水稻幼苗基因表現的組群(profile)。他們自發芽後五天的白化水稻幼苗萃取 RNA、轉換成雙股 cDNA，利用 SAGE 的分析技術，得到 10,122 個標記，其中 5,593 個標記出現的次數超過一次以上，最後總共得到 5,921 個不同的標記。然後將這些標記利用 SAGE 軟體依照其出現次數做一排列，並進行資料庫的比對，發現表現量大(出現次數多)的前 100 個標記中，80 %與已知的水稻 cDNA

或 EST 序列的相似度很高，此情況與在人類和酵母菌的 SAGE 分析結果相符。而進一步分析這 100 個大量表現的標記發現，其和許多的家管基因(housekeeping genes)，如參與醣類、核酸及胺基酸代謝的基因，具有很高的相似度，推測這些基因的表現是維持植物細胞生命力與個體生長所不可或缺。除此之外，在此高表現量基因中尚有一群與受逆境誘導的基因具有高相似度，而目前對於這些基因的功能仍然未知，值得進一步研究。

另外作者亦評估將 SAGE 技術用於選殖具差異性表現基因的可行性：利用經過缺氧處理(anaerobic treatment)一天的水稻與未處理的水稻當對照，進行 SAGE 分析；在缺氧處理的水稻上獲得 2,094 個標記，經分析後得到 1,496 個不同的標記；而未處理的對照，則得到 2,205 個標記，二處理經過比較後，得到 1,551 個差異的標記。將兩者的標記相互比較發現，絕大部分的標記表現量(於整個族群出現次數的比例)並無差異。而在 24 個具有差異表現的標記中，8 個基因係缺氧逆境所誘導的，而有 6 個基因在缺氧逆境下表現量會受到抑制。值得注意的是，水稻利用 SAGE 方法所篩選的基因有極大部分(10%)是功能未知的基因，或可能為新的基因(20%)(即在資料庫中找不到相似序列)(見圖三)。所以如能利用這些標記選殖相對的基因並加以研究，應可能解開植物抗缺氧逆境的機制。

從這些結果可知，將 SAGE 技術應用於植物上，可獲得許多未知的基因表現資訊，如能將這些基因予以選殖並進行功能分析，將有助於我們對植物基因表現與基因功能的瞭解。



圖三、利用 SAGE 技術分析水稻幼苗基因轉錄組群所得 100 個高度表現基因及其可能功能與頻度分佈圖。

另外為了加快基因選殖的腳步，作者捨棄傳統 SAGE 的做法，利用 RACE (rapid amplified cDNA end)技術，即利用此 9-13 個鹼基標記的序列當引子，配

合載體本身具有的序列引子(如 M13 forward 及 T7 promoter)，以整個基因庫的噬菌體 DNA 當模板，進行 PCR 擴增，進而選殖此產物而得到全長 cDNA，取代以 9-13 個鹼基的標記當探針進行基因庫篩選，這使我們更容易選殖到基因，也省去篩選基因庫的煩繁瑣步驟。

SAGE 分析法之優點：1.可同時分析所有的轉錄訊息，不需事先篩選任何已知資料。2.可針對已知及未知基因提供基因表現質與量的資訊。3.最適合用於探討或選殖二具差異的基因表現處理(如無氧或有氧處理)或材料(如正常細胞與癌細胞)之差異表現基因。

SAGE 分析法最主要的缺點是，必須使用大量的 mRNA (2.5-5μ g)進行試驗。這對偵測少量或微量樣品的基因表現是一極大的限制。其次，SAGE 分析過程所需繁瑣的純化步驟都會造成大量 mRNA 樣品的流失。MicroSAGE 法(Datson et al., 1999)即是針對 SAGE 的缺點改良而成的，其所有反應均在同一試管內進行，並以有限的 PCR 循環擴增足夠的雙標記(ditag)基因轉錄產物。因此，利用改良的 MicroSAGE 分析法使對於微量組織或特定組織之基因表現分析成為可行。

總之 SAGE 技術不僅可應用於動物，在植物上的應用亦將越來越多，隨著許多模式植物基因組分析的快速進展，建立各種不同環境、組織差異表現基因的資料庫已刻不容緩，研究基因表現使我們得知基因的功能，才能進一步利用基因。SAGE 技術提供了一快速的分析方法，有助於表現基因資料庫的快速建立。今後，將基因組序列資料庫與表現基因資料庫相配合，我們才能更加瞭解基因，更能發揮並利用基因的功能。

參考文獻：

1. **Datson, N. A., van der Perk-de Jong, J., van der Berg, M. P., de Kloet, E. R., and Vreugdenhil, E.** 1999. MicroSAGE: a modified procedure for serial analysis of gene expression in limited amounts of tissue. Nucl. Acid. Res. 27: 1300-1307.
2. **Matsumura, H., Nirasawa, S., and Terauchi, R.** 1999. Transcript profiling in rice (*Oryza sativa* L.) seedlings using serial analysis of gene expression (SAGE). Plant J. 20: 719-726.
3. **Velculescu, V. E., Zhang, L., Vogelstein, B., and Kinzler, K. W.** 1995. Serial

analysis of gene expression. Science 270: 484-487.

4. **Zhang, L., Zhou, W., Velculescu, V. E., Kern, S. E., Hruban, R. H., Hamilton, S. R., Volgelstein, Kinzler, K. W.** 1997. Gene expression profiles in normal and cancer cells. Science 276:1268-1272.

5. **SAGE 網站：**http://www.sagenet.org/**，可免費取得操作手冊及分析軟體。**SAGE **分析試劑組在** Genzyme Molecular Oncology (GMO)**公司已有商品可購買，此外另有利用** PCR **技術結合** SAGE **所發展的快速基因表現分析法** RAGE (rapid analysis of gene expression)**。**